



Universidad
Carlos III de Madrid



This is a postprint version of the following published document:

Grané, A. y Veiga, H. (2014). Outliers, GARCH-type models and risk measures: a comparison of several approaches. Journal of Empirical Finance, v. 26, pp. 26-40. Available in: <http://dx.doi.org/10.1016/j.jempfin.2014.01.005>

© Elsevier



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Outliers, GARCH-type models and risk measures: A comparison of several approaches[☆]

Aurea Grané^{a,*}, Helena Veiga^{a,b}

^a *Statistics Department, Universidad Carlos III de Madrid, C/Madrid, 126, 28903 Getafe, Spain*

^b *Instituto Flores de Lemus (Universidad Carlos III de Madrid), Finance Research Centre-Unide, Avenida das Forças Armadas, 1600-083 Lisboa, Portugal*

A B S T R A C T

In this paper we focus on the impact of additive outliers (level and volatility) on the calculation of risk measures, such as minimum capital risk requirements. Through simulation and empirical studies, we compare six alternative proposals that are used in the literature to reduce the effects of outliers in the estimation of risk measures when using GARCH type models. The methods are based on [1] correcting for significant outliers, [2] accommodating outliers using complex (e.g. fat tail) distributions and [3] accounting for outlier effects by robust estimation. The main conclusions of the simulation study are that the presence of outliers bias these risk measures, being the proposal by Grané and Veiga (2010) that providing the highest bias reduction. From the out of sample results for four international stock market indexes we found weak evidence that more complex models (specification and error distribution) perform better in estimating the minimum capital risk requirements during the last global financial crisis.

1. Introduction

The increase in volatility over recent years and the cataclysm involving financial markets across the world, especially since September 2008, have created an urgent need to protect the finance and banking system against large trading losses. After the Basel Accord of 1988, the first measure to tackle the problem was taken by demanding that the financial institutions reserve part of their capital to absorb a pre specified percentage of these unforeseen losses, a measure known as minimum capital risk requirements (MCRRs) (see for example Hsieh, 1993; Brooks et al., 2000). This measure is very similar to the well known value at risk (VaR) but differs from that in the following sense: a MCRR equal to p means that it is necessary to reserve a $p\%$ of the initial capital to cover a pre specified percentage of unforeseen losses.

Following the 1995 amendment of the Basel Accord, banks were allowed to use internal models to calculate the thresholds of their risk measures. This amendment was an attempt to remedy the fact that the standard approach to the estimation of the minimum capital risk requirements led to very conservative estimates and, consequently, to a waste of valuable resources by financial institutions that used the standard approach. Nevertheless, the recent poor evolution of financial markets underlines the low protection of financial institutions against extreme events, as well as the importance of forecasting volatility accurately for providing good estimates of these risk measures.

The accurate estimation of minimum capital risk requirements depends crucially on the accuracy of parameter estimates and volatility forecasts. Several GARCH type models have been proposed in the literature to capture the main features of financial

[☆] Financial support from research grants MTM2010-17323, ECO2009-08100 and ECO2012-32401 (Spanish Ministries of Science and Innovation and Economy and Competitiveness).

* Corresponding author.

E-mail addresses: aurea.grane@uc3m.es (A. Grané), mhveiga@est-econ.uc3m.es (H. Veiga).

time series and forecast volatility, due to their easy applicability and effectiveness in parameterizing the higher order dependence. However, it has been observed that the estimated residuals from this type of models still register excess kurtosis (see [Baillie and Bollerslev, 1989](#); [Teräsvirta, 1996](#)). One possible reason for this occurrence is that some observations on returns are not fitted by a Gaussian GARCH model, and not even by a t distributed GARCH model. These observations may be influential (see [Zhang, 2004](#) for a detailed definition of influential observation) since they can undesirably affect the estimation of parameters (see for example [Fox, 1972](#); [Van Dijk et al., 1999](#); [Verhoeven and McAleer, 2000](#)), the conditional homoscedasticity tests (see [Carnero et al., 2007](#); [Grossi and Laurini, 2009](#)) and the out of sample volatility forecasts (see for instance [Ledolter, 1989](#); [Chen and Liu, 1993a](#); [Franses and Ghijssels, 1999](#); [Carnero et al., 2012](#)). When this is the case, some authors denote such observations by outliers and distinguish between additive and innovational (or innovations) outliers.

This paper focuses mainly on the study of the effects of additive outliers (level and volatility) on the estimation of MCRRs for short and long trading investment positions. The effects of innovational outliers on the dynamic properties of the series are less important because they are propagated by the same dynamics, as in the rest of the series (see for example [Peña, 2001](#)). By means of an intensive simulation experiment, we study the outlier effects on the estimation of MCRRs and compare six alternative proposals. In particular, the methods under evaluation can be classified in three different categories: [1] correcting for significant outliers, [2] accommodating outliers using complex distributions and [3] accounting for outlier effects by robust estimation. The first approach is motivated by the fact that it generates volatility estimates with the smallest mean square error (see [Carnero et al., 2008](#)) and, therefore, potentially more accurate MCRRs. This approach includes the proposal by [Grané and Veiga \(2010\)](#) with hard and soft thresholding, which proceeds by detecting and correcting outliers before estimating these risk measures with symmetric and asymmetric GARCH type models. In the same category, we analyze the proposal by [Franses and Ghijssels \(1999\)](#) only for the GARCH model. The second approach consists in fitting, respectively, t distributed or skewed t distributed GARCH type models directly to the data. Finally, the third approach considers the robust estimation by [Muler and Yohai \(2008\)](#). These methods are exhaustively tested for out of sample conditional coverage, whenever possible.

Our study is informative, impartial and oriented to practitioners that want to use GARCH type models in the calculation of risk measures. The use of outlier detection and correction is justified by the fact that extreme events in financial markets are often not predictable. However, we need a reliable and if possible easy to apply method or model that performs the most accurate in the out of sample period.

The most important findings in this paper are: Firstly, outliers affect seriously the estimates of MCRRs and the effects depend on outlier magnitudes. It is often the case that the larger the outlier magnitude the larger the biases. Secondly, the detection proposal by [Grané and Veiga \(2010\)](#) with hard thresholding correction almost eliminates the biases on the MCRR estimates, whereas the highest biases are obtained when applying [Muler and Yohai's \(2008\)](#) proposal or, in the case of asymmetric models, when considering more complex error distributions. Thirdly, the impact on MCRRs is extremely high in the presence of additive volatility outliers. Fourthly, the results from the study of model uncertainty support that fitting asymmetric GARCH type models to simulated series from a GARCH(1,1) leads to similar MCRRs (the contrary statement does not hold). Finally, the empirical application and the out of sample results for four international stock market indexes seem to indicate that all the methods under study provide statistically equivalent failure rates in estimating the MCRRs for a 95% coverage during the last global financial crisis. However, using both more general models and more complex distributions generate failure rates closer to 5%.

The organization of this paper is as follows: In [Section 2](#) we describe how additive outliers are placed in the GARCH type models used in this paper. In [Section 3](#) we present the detection and correction methods by [Grané and Veiga \(2010\)](#), [Franses and Ghijssels \(1999\)](#) and the robust estimation by [Muler and Yohai \(2008\)](#). In [Section 4](#) we perform a simulation study in order to evaluate the effects of outliers on the estimation of MCRRs, and compare six different proposals. Complete results of the simulation study can be found in <http://www.est.uc3m.es/agrane/eng/public.html>. Additionally, we analyze the impact of model uncertainty on the estimation of MCRRs. In [Section 5](#) we test out of sample the presented proposals on four daily stock market indexes and we conclude in [Section 6](#).

2. Additive outliers in GARCH-type models

In this work we analyze the impact of additive outliers in symmetric and asymmetric GARCH type models. We start this Section by presenting the models under study. Next, we specify how additive outliers are included in these models. In particular, we consider two categories of additive outliers: additive level outliers (ALOs), which exert an effect on the level of the series but not on the evolution of the underlying volatility, and additive volatility outliers (AVOs), which also affect the conditional variance (see [Hotta and Tsay, 1998](#); [Sakata and White, 1998](#)).

2.1. GARCH type models

Return series of financial assets, although serially uncorrelated, are not independent because they contain higher order dependence. One way of parameterizing this dependence is by using models of autoregressive conditional heteroscedasticity such as the GJR(1,1) by [Glosten et al. \(1993\)](#), that is given by:

$$\begin{aligned} y_t &= \mu + \varepsilon_t, \\ \sigma_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1}^2 I_{\{\varepsilon_{t-1} < 0\}}(t-1) + \beta_1 \sigma_{t-1}^2, \end{aligned} \quad (1)$$

where μ is the conditional mean of the asset return y_t , $\varepsilon_t = \sigma_t \epsilon_t$ is the prediction error, $\sigma_t > 0$ is the conditional standard deviation of the underlying asset return (denoted volatility), $I_{\{\varepsilon_t < 0\}}(t) = 1$ if $t \in \{\varepsilon_t < 0\}$ and 0 otherwise and the error ϵ_t is iid $\sim N(0, 1)$. Furthermore, $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$ and $\gamma_1 \geq 0$ to guarantee a positive conditional variance, $\alpha_1 + \beta_1 + \gamma_1/2 < 1$ to enforce stationarity (Duan et al., 2006) and $\beta_1^2 + 2\beta_1\alpha_1 + 3\alpha_1^2 + \beta_1\gamma_1 + 3\alpha_1\gamma_1 + \frac{3}{2}\gamma_1^2 < 1$ to guarantee the existence of the fourth moment (Ling and McAleer, 2002).

One of the properties of this model is that it is able to accommodate the possibility that positive and negative price shocks affect the conditional variance σ_t^2 differently. Moreover, the model includes the GARCH(1,1) by Bollerslev (1986), that is obtained when $\gamma_1 = 0$. Finally, one extension of both models consists in allowing the error ϵ_t to follow a Student's t distribution (see Bollerslev, 1987 for the GARCH(1,1)).

2.2. Additive level outliers

These outliers can be caused by an institutional change or a market correction that does not affect volatility (see Doornik and Ooms, 2005). The conditional mean equation of the GJR(1,1) model with an ALO is defined as:

$$y_t = \mu + \omega_{AO} I_T(t) + \varepsilon_t,$$

where ε_t is defined as before, ω_{AO} represents the magnitude (or size) of the additive level outlier and $I_T(t) = 1$ for $t \in T$ and 0 otherwise, representing the presence of the outlier at a set of times T . In the study of the outliers' impact, we assume that the errors ϵ_t are iid normally or Student t or skewed t distributed. Eq. (1) of the conditional variance for the GJR(1,1) model remains the same, since this type of outliers only modifies the equation of the conditional mean of the asset return.

2.3. Additive volatility outliers

These outliers affect not only the volatility but also the series level. Their effect on the original series is similar to a patch of ALOs with decreasing magnitudes when $\beta_1 < 1$. In this context, an AVO for the GJR(1,1) model is defined as:

$$\begin{aligned} y_t &= \mu + \varepsilon_t^*, \\ \varepsilon_t^* &= \omega_{AO} I_T(t) + \varepsilon_t, \\ \sigma_t^{*2} &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^{*2} + \gamma_1 \varepsilon_{t-1}^{*2} I_{\{\varepsilon_{t-1} < 0\}}(t-1) + \beta_1 \sigma_{t-1}^{*2}, \end{aligned} \quad (2)$$

where $\varepsilon_t^* = \sigma_t^* \epsilon_t$ and $I_{\{\varepsilon_t < 0\}}(t) = 1$ if $t \in \{\varepsilon_t < 0\}$ and 0 otherwise. We can express σ_t^{*2} in terms of the dynamic effect of the outlier by replacing ε_t^* in Eq. (2):

$$\begin{aligned} \sigma_t^{*2} &= \alpha_0 + \left(\alpha_1 + \gamma_1 I_{\{\varepsilon_{t-1} < 0\}}(t-1) \right) \varepsilon_{t-1}^2 \\ &\quad + \left(\alpha_1 + \gamma_1 I_{\{\varepsilon_{t-1} < 0\}}(t-1) \right) \left(2 \omega_{AO} \varepsilon_{t-1} + \omega_{AO}^2 \right) I_T(t-1) + \beta_1 \sigma_{t-1}^{*2}. \end{aligned} \quad (3)$$

Finally, for the GARCH(1,1) model the AVO and its effects on σ_t^{*2} are obtained when $\gamma_1 = 0$ in Eqs. (2) and (3), respectively.

3. Procedures under study

In this Section we resume the outlier detection proposals by Grané and Veiga (2010), Franses and Ghijsels (1999) and the robust estimation by Muler and Yohai (2008). We refer the reader to the original papers for the details.

3.1. Wavelet based detection and correction procedures

Grané and Veiga (2010) proposed a general detection and correction method based on wavelets, which can be applied to a large class of volatility models. The effectiveness of their proposal was tested applying it to several volatility models, such as the GARCH(1,1), the GJR(1,1) and the autoregressive stochastic volatility model, ARSV(1) by Taylor (1986), with errors following a Gaussian or a Student's t distribution and comparing it with the proposals by Bilen and Huzurbazar (2002), Franses and Ghijsels (1999) and Doornik and Ooms (2005). The intensive Monte Carlo study revealed that Grané and Veiga's (2010) proposal is not only as good as these alternatives in detecting different type of outliers (isolated ALOs, multiple ALOs, AVOs and patches of ALOs), but it is also much more reliable, since it detects a significantly smaller number of false outliers.

The algorithm uses the notions of discrete wavelet transform and inverse discrete wavelet transform (see Percival and Walden, 2000 for a complete guide to wavelet methods for time series). In particular, the proposal is based on the wavelet detail coefficients resulting from the discrete wavelet transform of the series of residuals, which are obtained after fitting a particular volatility model. The outliers are identified as those observations in the original series whose wavelet detail coefficients are greater (in absolute value) than a certain threshold. (We refer the reader to the original paper for the details concerning the threshold computation).

Once the outlier positions in the series have been determined (using the series of residuals), we propose correcting those observations from the series of returns by hard thresholding or alternatively by soft thresholding. In [Grané and Veiga \(2010\)](#), we only considered correcting by hard thresholding, but since soft thresholding has become popular in the context of wavelet estimation and there is some evidence that for some particular situations it turns out to be superior to hard thresholding (see [Droge, 2006](#) for more details), we have included it for comparison.

3.2. Franses and Ghijssels's (1999) proposal

[Franses and Ghijssels \(1999\)](#) exploited the analogy of the GARCH(1,1) model with an ARMA(1,1) model to adapt [Chen and Liu's \(1993b\)](#) method to detect and correct additive level outliers in GARCH(1,1) models. In particular, the conditional variance equation of the GARCH(1,1) model (that is, Eq. (1) with $\gamma_1 = 0$) can be rewritten as an ARMA(1,1) for ε_t^2 :

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1)\varepsilon_{t-1}^2 + v_t - \beta_1 v_{t-1}, \quad (4)$$

where $v_t = \varepsilon_t^2 - \sigma_t^2$. From the previous equation, v_t can be written as $v_t = \frac{\alpha_0}{1 - \beta_1 L} + \pi(L)\varepsilon_t^2$ with $\pi(L) = \frac{1 - (\alpha_1 + \beta_1)L}{1 - \beta_1 L}$ and L is the lag operator. For convenience and without loss of generality, we suppose that the conditional mean equation now is $y_t = \sigma_t \varepsilon_t$. Instead of the true series ε_t , we observe e_t defined by $e_t^2 = \varepsilon_t^2 + \omega_{AO} I_T(t)$, where, as before, ω_{AO} represents the magnitude (or size) of the additive level outlier and $I_T(t) = 1$ for $t \in T$ and 0 otherwise, representing the presence of the outlier at a set of times T . When the ARMA model Eq. (4) is fitted to the observed e_t^2 , the residuals are

$$\eta_t = v_t + \pi(L)\omega_{AO} I_T(t). \quad (5)$$

Eq. (5) can be seen as a regression model of η_t on x_t , $\eta_t = \omega_{AO}x_t + v_t$, where

$$x_t = \begin{cases} 0, & \text{if } t < \tau, \\ 1, & \text{if } t = \tau, \\ -\pi_k, & \text{if } t > \tau + k \text{ and } k > 0. \end{cases}$$

We are assuming that there is an outlier of size ω_{AO} at time $t = \tau$. According to these authors, the detection of an ALO is based on the following test statistic:

$$\hat{\tau}(\tau) = \frac{\sum_{t=\tau}^n x_t \eta_t}{\hat{\sigma}_v \sqrt{\sum_{t=\tau}^n x_t^2}},$$

where $\hat{\sigma}_v$ is the estimated standard deviation of the residuals.¹ We refer the reader to the original paper by [Franses and Ghijssels \(1999\)](#) for details on the detection algorithm of an ALO (see also [Charles and Darné \(2005\)](#)). We must say that the good performance of this proposal depends heavily on the correct selection of the critical value ([Charles \(2008\)](#) chose a critical value of 10 based on the simulation results by [Verhoeven and McAleer \(2000\)](#) and [Franses and van Dijk \(2002\)](#)).

3.3. Muler and Yohai's (2008) robust estimation

In [Muler and Yohai \(2008\)](#) a class of robust estimates for GARCH models is proposed. They are called bounded M estimates (BM estimates) and are based on a modification of M estimates that include a mechanism that restricts the propagation of outlier effects on the subsequent predictors of the conditional variances. These authors proved their consistency and asymptotic normality and also their robustness in the sense of [Huber \(1981\)](#), that is, BM estimates have a high efficiency under a Gaussian GARCH model and are not much influenced by a small fraction of outlying observations.

Formally, given the parameter values $\mathbf{c} = (\mathbf{a}, \mathbf{b})$, where $\mathbf{a} = (a_0, a_1)$, $\mathbf{b} = b_1$, these authors define for all t

$$\sigma_{t,k}^{2*}(\mathbf{c}) = a_0 + a_1 \sigma_{t-1}^{2*}(\mathbf{c}) r_k \left(\frac{\varepsilon_{t-1}^2}{\sigma_{t-1,k}^{2*}(\mathbf{c})} \right) + b_1 \sigma_{t-1,k}^{2*}(\mathbf{c}),$$

where $\varepsilon_t = 0$ for $t \leq 0$ and

$$r_k(u) = \begin{cases} u & \text{if } u \leq k, \\ k & \text{if } u > k. \end{cases}$$

The propagation of the effect of one outlier in time t on the conditional variance $\sigma_{t',k}^{2*}(\mathbf{c})$, $t' > t$, disappears after a few periods. Therefore, if ε_t follows a GARCH(1,1) model with some outliers, the M estimates using these conditional variances would fit well the data.

¹ See [Franses and Ghijssels \(1999\)](#) for details on the estimation of the standard deviation.

Table 1

M0 method. MCRRs for 95% coverage probability as a percentage of the initial value of the simulated series (standard deviation) and e_r stands for the relative error.

			GARCH(1,1)				GJR(1,1)			
			Long position		Short position		Long position		Short position	
		n	MCRR	e_r	MCRR	e_r	MCRR	e_r	MCRR	e_r
s1	1 ALO	500	1.629 (0.583)	0.037	1.659 (0.625)	0.038	2.188 (1.058)	0.044	2.240 (1.076)	0.048
	of size	1000	1.591 (0.524)	0.020	1.622 (0.547)	0.021	2.174 (1.100)	0.028	2.231 (1.198)	0.030
	$\omega_{AO} = 5\sigma_y$	5000	1.576 (0.510)	0.004	1.605 (0.532)	0.004	2.109 (0.903)	0.005	2.163 (0.959)	0.004
	1 ALO	500	1.746 (0.790)	0.111	1.783 (0.857)	0.116	2.512 (3.183)	0.198	3.331 (26.212)	0.558
	of size	1000	1.673 (0.580)	0.072	1.708 (0.614)	0.075	2.307 (1.351)	0.091	2.377 (1.521)	0.098
	$\omega_{AO} = 10\sigma_y$	5000	1.595 (0.529)	0.017	1.624 (0.555)	0.016	2.135 (0.917)	0.017	2.191 (0.976)	0.017
	1 ALO	500	1.863 (1.665)	0.194	1.935 (2.229)	0.211	2.766 (2.812)	0.320	2.947 (3.939)	0.378
	of size	1000	1.788 (0.827)	0.146	1.831 (0.896)	0.152	2.491 (1.601)	0.178	2.583 (1.885)	0.193
	$\omega_{AO} = 15\sigma_y$	5000	1.625 (0.569)	0.036	1.656 (0.604)	0.036	2.185 (1.000)	0.041	2.244 (1.081)	0.042
	2 ALOs	500	2.080 (1.599)	0.324	2.151 (1.838)	0.346	2.999 (2.820)	0.431	3.191 (3.821)	0.493
s2	of size	1000	1.920 (0.994)	0.231	1.975 (1.137)	0.243	2.799 (1.936)	0.324	2.922 (2.257)	0.350
	$\omega_{AO} = 15\sigma_y$	5000	1.694 (0.611)	0.080	1.729 (0.646)	0.082	2.289 (1.078)	0.091	2.356 (1.175)	0.094
	Patch 3 ALOs	500	2.117 (1.494)	0.348	2.191 (1.760)	0.373	2.930 (2.747)	0.398	3.111 (3.466)	0.455
s3	of size	1000	1.825 (0.787)	0.170	1.868 (0.855)	0.176	2.503 (1.772)	0.184	2.600 (2.174)	0.200
	$\omega_{AO} = 15\sigma_y$	5000	1.642 (0.706)	0.047	1.676 (0.779)	0.049	2.196 (1.086)	0.046	2.260 (1.212)	0.049
	1 AVO	500	8.161 (7.326)	4.195	10.231 (12.459)	5.402	3.345 (2.930)	0.596	3.618 (3.539)	0.692
s4	of size	1000	5.541 (6.799)	2.552	6.855 (11.289)	3.314	2.664 (1.782)	0.260	2.807 (2.023)	0.297
	$\omega_{AO} = 15\sigma_y$	5000	2.732 (3.887)	0.741	3.081 (6.126)	0.928	2.182 (1.044)	0.040	2.252 (1.138)	0.045
	No outliers	500	1.571 (0.579)		1.598 (0.620)		2.096 (1.004)		2.138 (1.011)	
s5a	(fitting the same model)	1000	1.560 (0.530)		1.589 (0.551)		2.114 (0.993)		2.165 (1.047)	
		5000	1.569 (0.506)		1.598 (0.526)		2.099 (0.904)		2.154 (0.960)	
s5b	No outliers	500	1.566 (0.581)		1.593 (0.624)		2.116 (1.002)		2.164 (1.011)	
	(fitting the other model)	1000	1.559 (0.541)		1.589 (0.563)		2.115 (0.950)		2.167 (0.999)	
		5000	1.569 (0.508)		1.598 (0.529)		2.106 (0.864)		2.161 (0.914)	

GARCH(1,1) model with parameter values $\{\alpha_0 = 0.01, \alpha_1 = 0.08, \beta_1 = 0.91\}$ and GJR(1,1) model with parameter values $\{\alpha_0 = 0.02, \alpha_1 = 0.03, \beta_1 = 0.91, \gamma_1 = 0.10\}$.

4. Simulation study

In this Section we study two situations. In the first one, we consider that there is no uncertainty about the data generating process (DGP), whereas in the second one, we analyze the impact of uncertainty on the MCRRs when the data generating process is unknown. In this case, we consider two sources of uncertainty: functional form and error distribution.

In this study we consider two Gaussian models as data generating processes: a symmetric volatility model, GARCH(1,1) and an asymmetric one, GJR(1,1). Parameter values are specified in Table 1. We start by simulating return series of different sample sizes ($n = 500, 1000, 5000$) from the considered model with parameter values chosen to resemble those obtained for time series of financial returns.² The frequency of the simulations is daily. The outliers are placed randomly in the series and each scenario (given n and ω_{AO}) involves 1000 Monte Carlo samples. Next we describe the considered situations:

- s1 One isolated ALO of three different magnitudes ($\omega_{AO} = 5\sigma_y$, $\omega_{AO} = 10\sigma_y$, $\omega_{AO} = 15\sigma_y$) in simulated series from the considered model. For each magnitude, the sample sizes considered are $n = 500, 1000, 5000$ and σ_y is the standard deviation of the returns.
- s2 Two isolated ALOs of size $\omega_{AO} = 15\sigma_y$ in simulated series from the considered model of sample sizes of $n = 500, 1000, 5000$.
- s3 Patches of three ALOs of size $\omega_{AO} = 15\sigma_y$ in series simulated from the considered model of samples sizes of $n = 500, 1000, 5000$. The beginning of the patch is placed randomly in the series.
- s4 One isolated AVO of magnitude $\omega_{AO} = 15\sigma_y$ in series simulated from the considered model of samples sizes of $n = 500, 1000, 5000$.
- s5a Original simulated series (with no outliers) of sample sizes $n = 500, 1000, 5000$. The MCRRs are calculated with the same model used to simulate the series.
- s5b Original simulated series (with no outliers) of sample sizes $n = 500, 1000, 5000$. The MCRRs are calculated with a different model from that used to simulate the series. This situation will help to analyze the impact of model uncertainty (see Section 4.2).

Capital risk requirements, given by the percentage of the initial value of the position for 95% coverage, are estimated for a 1 day investment horizon for the simulated data. Therefore, we calculate the MCRRs for each Monte Carlo sample in each scenario

² Following an anonymous referee's suggestions, we have considered other parameter values in the simulations and obtained similar conclusions. The results are available from the authors upon request.

in the following way (this procedure is hereafter denoted as M0 method): We start by fitting the corresponding model to the simulated series, using the G@RCH 6.0 package by [Laurent and Peters \(2006\)](#), and we generate 20,000 paths of future values of the price series with the help of the parameter estimates, the disturbances obtained by sampling with replacement from the iid residuals (iid bootstrap), and the one day ahead volatility forecasts. The maximum loss over a given holding period, supposing there is only one futures contract is then obtained by computing $Q = (P_0 - P_1)$, where P_0 is the initial value of the position and P_1 is the lowest simulated price (for a long position) or the highest simulated price (for a short position) over the period. We assume that the position is open on the final day of the sample (see [Brooks, 2002; Brooks et al., 2000](#)). We can write $\frac{Q}{P_0} = (1 - \frac{P_1}{P_0})$ for a long position, and $\frac{Q}{P_0} = (\frac{P_1}{P_0} - 1)$ for a short position. Regarding that P_0 is constant, the distribution of Q only depends on the distribution of P_1 . Assuming that simulated prices are lognormally distributed, then the maximum loss for a long position over the simulated days is given by

$$Q/P_0 = 1 - \exp(c_\alpha s + m), \quad (6)$$

where s and m are the standard deviation and mean of the $\ln(P_1/P_0)$, respectively and c_α is the $\alpha \times 100$ th percentile of the standard normal distribution. Analogously, for a short position $Q/P_0 = \exp(c_1 - \alpha s + m) - 1$, where $c_1 - \alpha$ is the $(1 - \alpha) \times 100$ th percentile of the considered distribution. In general, the $\alpha \times 100$ th and $(1 - \alpha) \times 100$ th percentiles depend on the assumption assumed for the error distribution of the model.³ Finally, for each situation s1–s5b, MCRR estimates are obtained as the mean of 1000 estimated MCRR values. [Table 1](#) contains these results.

4.1. No uncertainty about DGP

We start by studying the effect of outliers on the MCRR estimates and next we compare six different proposals for reducing the MCRR estimation biases.

In [Table 1](#) we observe that: First, in general, concerning long position, the relative errors increase with the outlier magnitude and decrease with the sample size. Second, the relative error substantially increases in the presence of multiple isolated ALOs, even for large sample sizes ($n = 5000$). Third, patches of three ALOs induce higher biases than multiple isolated ALOs (of the same magnitude) for sample sizes equal to or greater than $n = 1000$. Fourth, the effect of AVOs on the MCRRs is the greatest, being the GJR model the most robust. Finally, similar effects are observed for short position, although with slightly greater relative errors.

[Fig. 1](#) suggests that these changes in the MCRRs may be due to the parameter estimate biases and volatility forecast biases when we include an outlier of magnitude $15\sigma_y$. In particular, panel (d) shows the changes in the distribution (box plot) of the one day ahead volatility forecasts when an outlier of magnitude $15\sigma_y$ is randomly included in each simulated series. We observe that in the absence of outliers the estimation of the volatility forecasts is more accurate (notice the less dispersion in the distributions and the low values for the estimates).

Given the evidence of the previous simulation results, pertinent questions are: Are the MCRR estimates obtained with Gaussian GARCH(1,1) and GJR(1,1) models similar to those obtained with more appropriate distributions in the presence of outliers, such as Student t or skewed Student t ? Are the former MCRR estimates similar to those obtained by robust GARCH estimation?

The second part of this Section deals with these issues and consists in the comparison of six proposals that can be classified in three approaches to deal with outliers when estimating MCRRs: [1] correcting for significant outliers, [2] accommodating outliers using complex distributions and [3] accounting for outlier effects by robust estimation. For all these methods, the calculation of the MCRRs is analogous to the M0 method, but instead of starting by fitting either a GARCH(1,1) or a GJR(1,1) directly to the simulated outlier contaminated series, we start by:

- M1 Filtering using the procedure by [Grané and Veiga \(2010\)](#) with hard thresholding and then fitting the considered model, [1] Correcting outliers.
- M2 Filtering using the procedure by [Grané and Veiga \(2010\)](#) with soft thresholding and next fitting the considered model,
- M3 Filtering using the procedure by [Franses and Ghijsels \(1999\)](#) and then fitting a GARCH(1,1) model,
- M4 Fitting t distributed GARCH(1,1) or GJR(1,1) models, with endogenous degrees of freedom, [2] Complex distributions.
- M5 Fitting skewed t distributed GARCH(1,1) or GJR(1,1) models, with endogenous degrees of freedom,
- M6 Fitting a Gaussian GARCH(1,1) using [Muler and Yohai's \(2008\)](#) robust BM2 estimation method. [3] Robust estimation.

In M4 and M5 methods, the c_α quantile of formula (6) is taken as $\alpha \times 100$ th percentile of the Student t distribution with v degrees of freedom and the $\alpha \times 100$ th percentile of the skewed Student t distribution with v degrees of freedom and asymmetry coefficient ξ , respectively. Analogously, for the $c_1 - \alpha$ quantile.

As a general comment on these methods, we can say that methods M1–M3 consist of detecting and correcting outliers before estimating the MCRRs, therefore in these cases MCRRs are computed from the corrected series. Concerning tbM3, the simulation

³ We consider short and long positions so that the conclusions may be applied to a wide variety of derivatives, such as futures, etc., whose losses are computed differently in case of selling (short position) or buying (long position) the derivative.

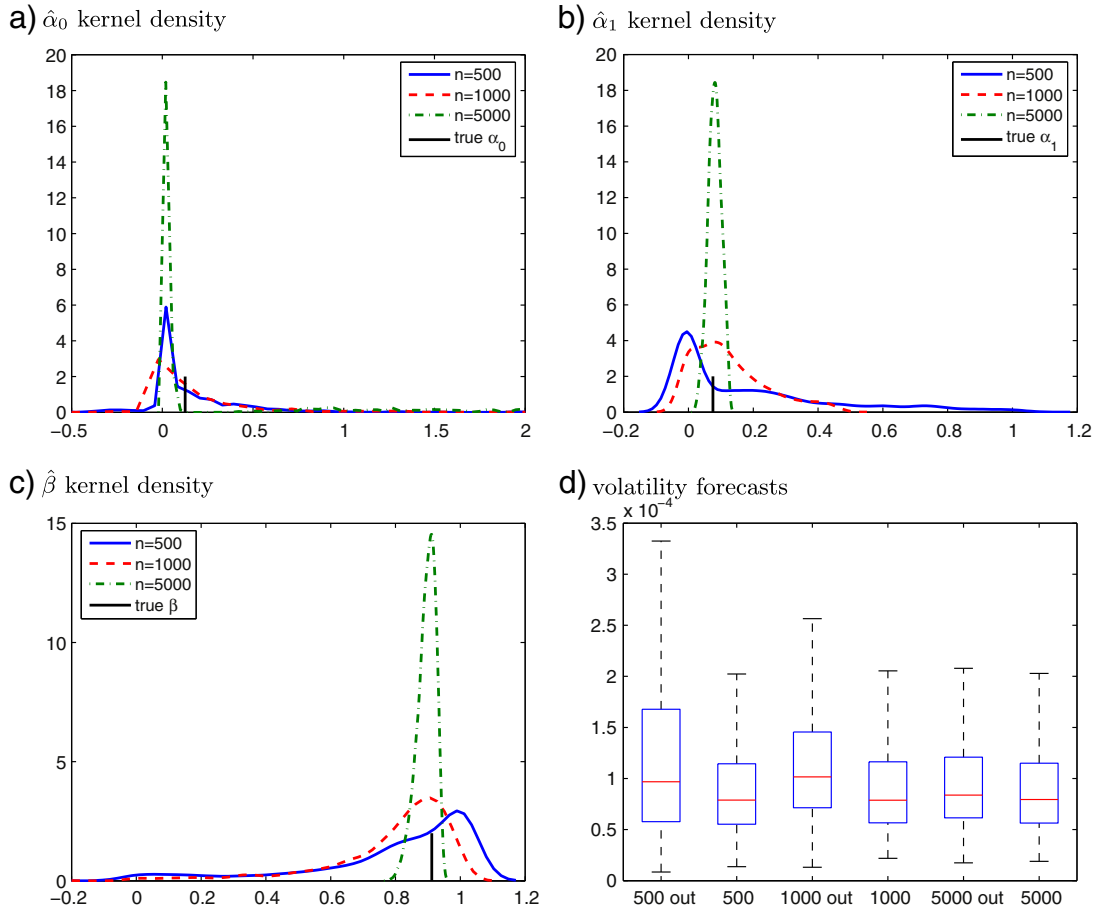


Fig. 1. Panels (a)–(c): Kernel densities of parameter estimators $\hat{\alpha}_0$, $\hat{\alpha}_1$ and $\hat{\beta}$ obtained from 1000 samples of size $n = 500, 1000, 5000$ from a Gaussian GARCH(1,1) with true parameter values $\alpha_0 = 0.12$, $\alpha_1 = 0.08$ and $\beta = 0.91$. One outlier of size $\omega_{AO} = 15\sigma_y$ is included randomly in each series. Panel (d): Box-plots for one-day-ahead volatility forecasts.

study was conducted only for $n = 500, 1000$ because we found no critical values for $n = 5000$ in the literature. Other models have not been considered in this case, because the procedure by Franes and Ghijssels (1999) was designed for detecting ALOs in GARCH models. Methods M4–M6 are alternative ways of dealing with outliers that consist in fitting models that can better accommodate these observations. In fact, M4 considers a Student t distribution for the model errors and M5 a skewed Student t distribution. Finally, M6 method consists of a GARCH robust estimation proposed by Muler and Yohai (2008) that only considers isolated ALOs in GARCH models (see Park (2002) and Carnero et al. (2012) for alternative robust techniques). The performances of all these methods in situations s1–s5 are summarized in Figs. 2–3. Complete results are available at <http://www.est.uc3m.es/agrane/eng/public.html>.

Fig. 2 summarizes situations s1–s4 concerning the relative errors in estimating the MCRRs using GARCH and GJR models for long position, and Fig. 3 for short position. In panels (a) of Figs. 2 and 3 we observe that approach [1] ‘correcting outliers’ (M1 and M3 methods) provides the smallest relative errors while using the GARCH model, whereas approaches [2] ‘complex distributions’ and [3] ‘robust estimation’ produce the highest relative errors, indicating that they tend to overestimate the MCRRs in the presence of outliers. Similar conclusions are derived from panels (c) of both figures, where the smallest biases are obtained by approach [1] (methods M1 and M2). However, we observe that the impacts on the MCRRs are extremely high in the case of AVOs.

Regarding GJR model, in panels (b) of Figs. 2 and 3 we observe that approach [1] ‘correcting for outliers’ (methods M1 and M2) is that with the smallest bias. Furthermore, there are almost no differences between the performance of approach [2] ‘complex distributions’ and M0, indicating that methods M4–M5 tend to overestimate the MCRRs in the presence of outliers. Similar conclusions are derived from panel (d), where the smallest biases are obtained by approach [1] (methods M1 and M2). However, the impact of AVOs on the MCRRs when using the GJR model is considerably smaller than that obtained when using the GARCH model. This leads us to conclude that the GJR model is more robust in the presence of AVOs.

Overall, we can say that applying models that are known to be more robust to the effect of outliers (approaches [2] and [3]) leads to an increase of the MCRRs estimates. This suggests that the capital reserved to absorb unforeseen losses should be greater than that estimated using Gaussian GARCH and GJR models. On the other hand, approach [1] (mainly M1 or M2 method)

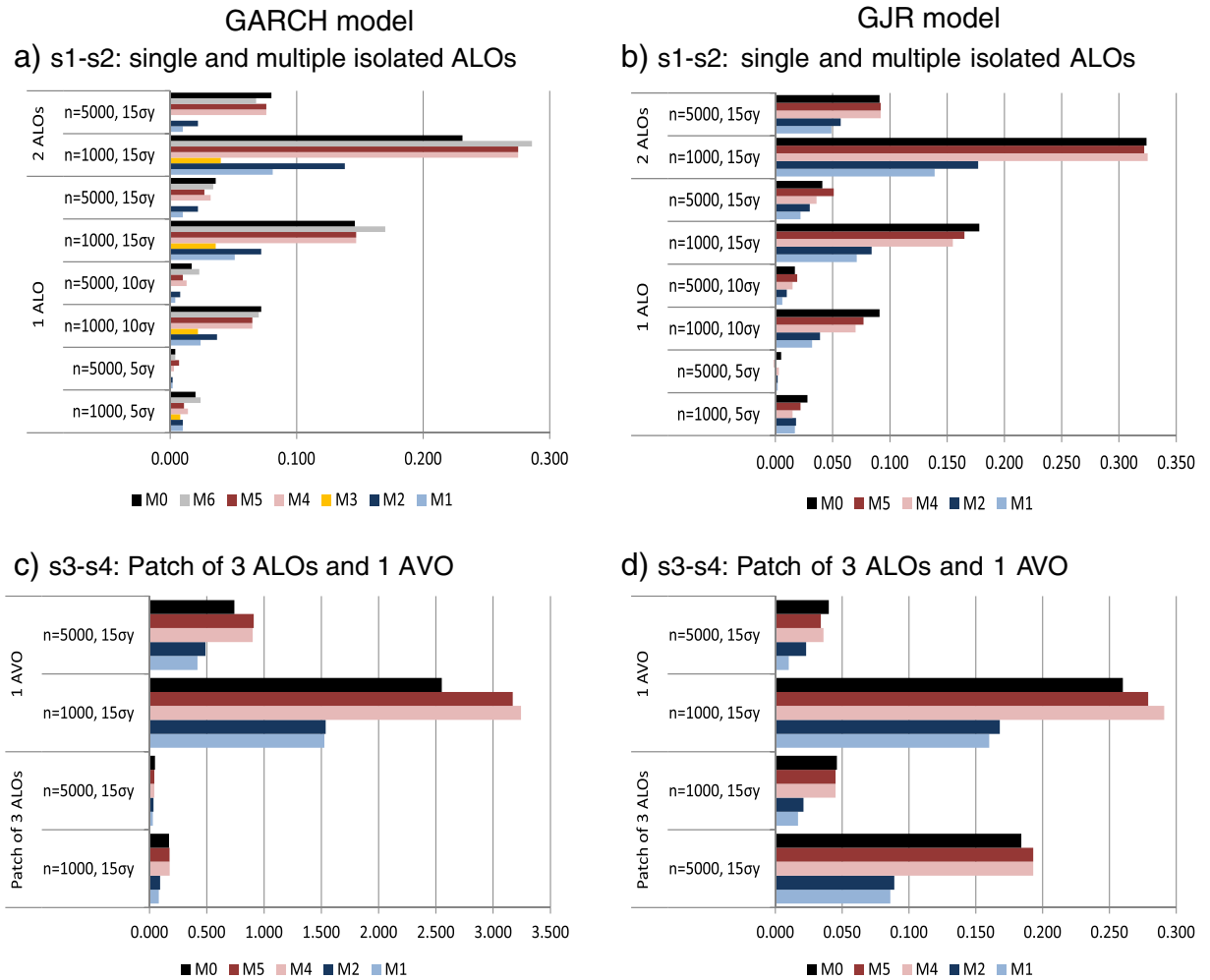


Fig. 2. Comparison of relative errors for long position.

generates MCRRs more similar to those obtained with simulated series that are not contaminated by the presence of outliers, especially when the sample size is large ($n = 500$). This means, that after correcting the series for outliers with methods M1 and M2, we are able to reduce the outlier effects on the computation of the MCRRs. Finally, we notice that the hard thresholding correction generates MCRRs closer to those of non contaminated series than the soft thresholding procedure.

4.2. Model uncertainty

In this Section we analyze the impact of uncertainty on the MCRRs when the simulated series are not contaminated with outliers. We consider two sources of uncertainty: functional form and error distribution. Complete results of the simulation study can be found in Table 1 and in <http://www.est.uc3m.es/agrane/eng/public.html>.

Panels (a) (b) of Fig. 4 contain a summary of the comparison of MCRRs concerning model specification and panels (c) (d) error distribution.⁴ From panel (a) we observe that fitting a Gaussian GJR model to simulated series from Gaussian GARCH leads to practically the same MCRRs for sample sizes equal or greater than $n = 100$. On the other hand, we obtain higher MCRRs for all sample sizes when a Gaussian GARCH model is fitted to simulated series from a Gaussian GJR (see panel (b)). This effect decreases with the sample size. A possible explanation is that since the GJR model includes asymmetry, when we fit the GARCH model, that does not include it, the GARCH overestimates volatility and consequently generates higher MCRRs, particularly for small sample sizes. Regarding error distribution uncertainty, we observe that fitting non Gaussian models to simulated series from Gaussian models (either GARCH or GJR) leads to smaller estimates of the MCRRs (see panels (c) and (d)). This may be due to

⁴ In each panel, the model used to generate the data is indicated by DGP and each graph contains the comparison of the estimation of the MCRRs using the DGP and the other considered models.

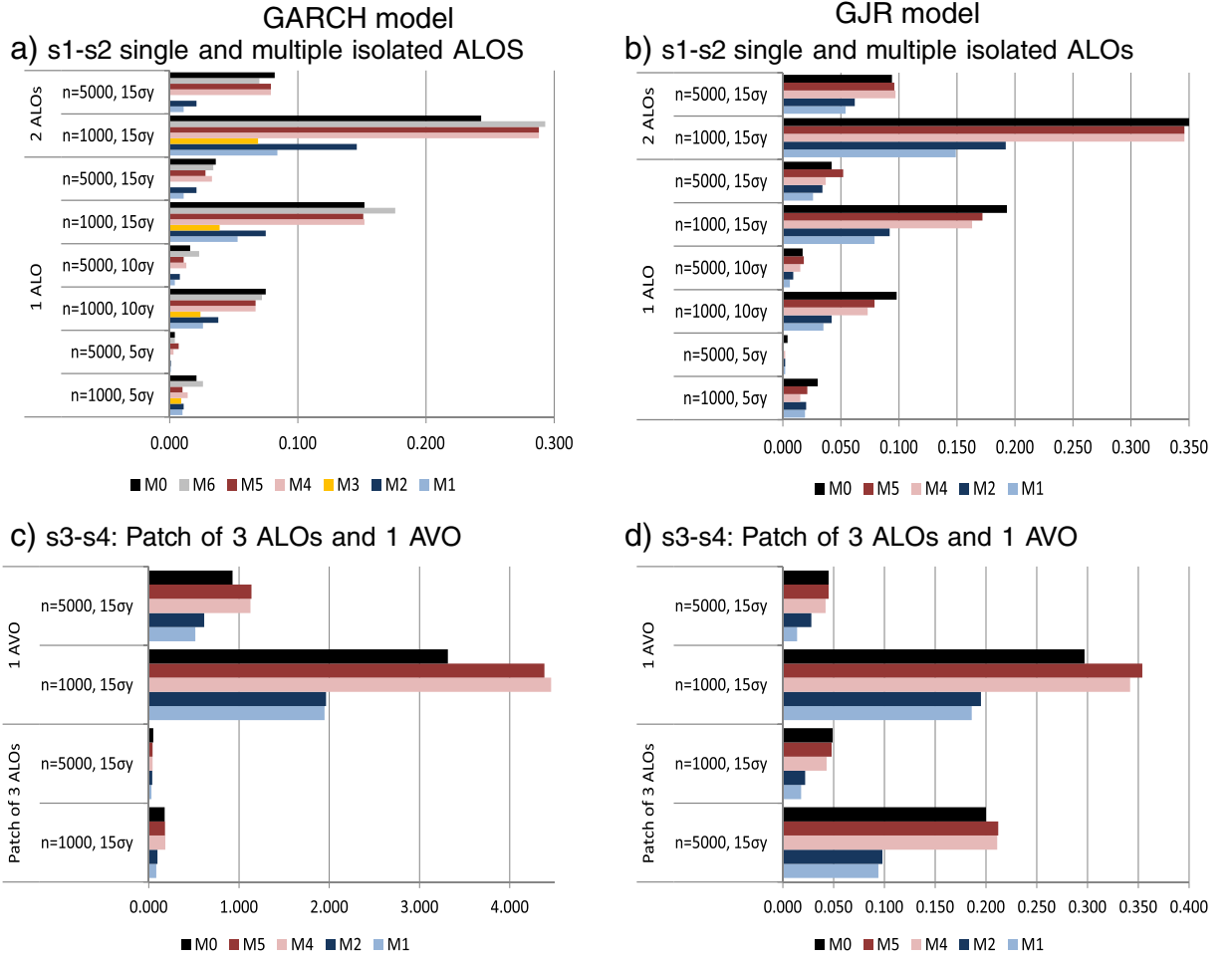


Fig. 3. Comparison of relative errors for short position.

the fact that non Gaussian GARCH models are much more flexible to model simultaneously kurtosis and volatility. Additionally, concerning the data generating process, we observe that the MCRRs are always higher when the simulated series come from a GJR model (notice the axis scale in panels (b) (d) versus (a) (c)).

Finally, we have seen that the MCRRs computed under situation s5b are smaller than those obtained in situation s5a (that is, when there is no uncertainty about the DGP). Since the biases on the MCRRs (see Figs. 2 3) were computed under situation s5a, we may expect that they would be higher under uncertainty. The only exception may be when a Gaussian GARCH is fitted to simulated series coming from Gaussian GJR.

5. Empirical applications

In this Section we evaluate the performance of the methods presented in Section 4.1, with series of real data. Note that, although these methods have been described for simulated data, we keep the same notation while working with real data.

We start by introducing four daily stock market indexes: the FTSE 100 index, the Nasdaq index, the NYSE composite index and the S&P 500 index. The data was collected from Yahoo Finance (<http://finance.yahoo.com/>) and spans the period of November 26, 1990 till February 19, 2013. In Fig. 5 we depict the four return series, $y_t = (\log p_t - \log p_{t-1}) \cdot 100$, where p_t is the value at time t of the corresponding index and in Table 2 we report some descriptive statistics.

From Fig. 5 we observe volatility clustering, periods of high volatility followed by periods of low volatility. On the other hand, the results of Engle's ARCH test (shown in Table 2) lead to the rejection of the null hypothesis of no conditional heteroskedasticity, which suggest that GARCH type models may be a good alternative to model the volatility of these financial return series. Additionally, the four return series are negatively skewed and have kurtosis ranging from 8.006 for the Nasdaq to 13.8450 for the NYSE index, characteristics of the data that can be reinforced due to the existence of outliers (see Charles, 2008).

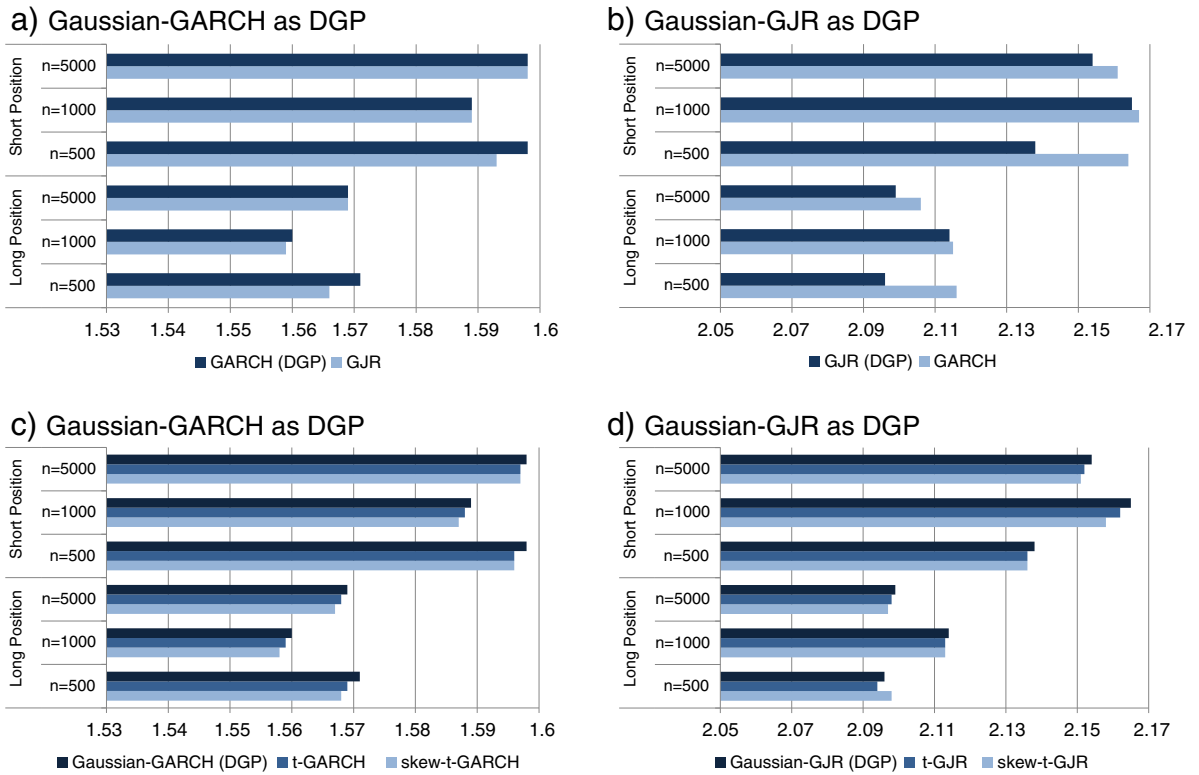


Fig. 4. Impact of uncertainty on the MCRRs. Panels (a)–(b) model specification and panels (c)–(d) error distribution.

In the GARCH model positive and negative unexpected changes in asset prices of the same magnitude generate the same amount of volatility. However, it is known that larger negative returns create more volatility than positive ones of the same magnitude. If this is the case, the GARCH model will underestimate the amount of volatility generated by negative returns (and overestimate the amount of volatility following positive returns). Moreover, if larger returns cause more volatility, this will also lead to the underestimation of volatility after a large shock (and overestimation of volatility after a small shock), affecting the estimates of risk measures. These findings suggest tests for sign and size biases. Following Engle and Ng (1993) we infer about the need of an asymmetric conditional heteroscedastic volatility model by estimating the following equation by OLS:

$$v_t^2 = \phi_0 + \phi_1 I_{t-1} + \phi_2 \varepsilon_{t-1} I_{t-1} + \phi_3 \varepsilon_{t-1} (1 - I_{t-1}) + e_t,$$

where $v_t^2 = (\varepsilon_t / \sigma_t)^2$, I_{t-1} is a dummy variable that takes value 1 if $\varepsilon_{t-1} < 0$ and 0 otherwise, and e_t is a white noise error. If the model is correctly specified then $\phi_1 = \phi_2 = \phi_3 = 0$ and e_t is iid. The significance of ϕ_1 indicates the presence of sign bias, that is, positive and negative realizations of ε_{t-1} affect future volatility differently. On the other hand, if ϕ_2 and ϕ_3 are statistically significant, this would suggest size bias. The model can be subject to all of these tests at once by testing $\phi_1 = \phi_2 = \phi_3 = 0$, which can be done, under Gaussianity, based upon the Lagrange Multiplier Principle with the statistic $n R^2$ that follows an χ^2 distribution with degrees of freedom equal to the number of restrictions under the null (n is the number of observations used in the regression and R^2 is the coefficient of determination). The results of this test are reported in Table 3. Overall, they suggest additional modeling structure that incorporates the possibility of asymmetry into the variance equation of the GARCH model.

5.1. Detection of outliers using wavelets

In order to check if there are outliers in the data, we apply Grané and Veiga's (2010) procedure to the residual series of the estimated GARCH(1,1) and GJR(1,1) models. Results are shown in Table 4. The same days are identified as possible outliers under either Gaussian GARCH(1,1) or GJR(1,1) model.⁵

From Table 4 we observe that some days are considered as outliers for more than one index, such as 11/15/91 and 02/27/07, that are detected as outliers in three return series. The detection results confirm the effectiveness of our procedure in capturing the most important crashes in the sample period for some important international stock markets. For instance, November 15,

⁵ The proposal by Franses and Ghijssels (1999) (method M3) is not used for checking the presence of outliers due to the unavailability of both critical values and studies that report the method's performance (under these critical values) for sample sizes comparable to those considered in this study.

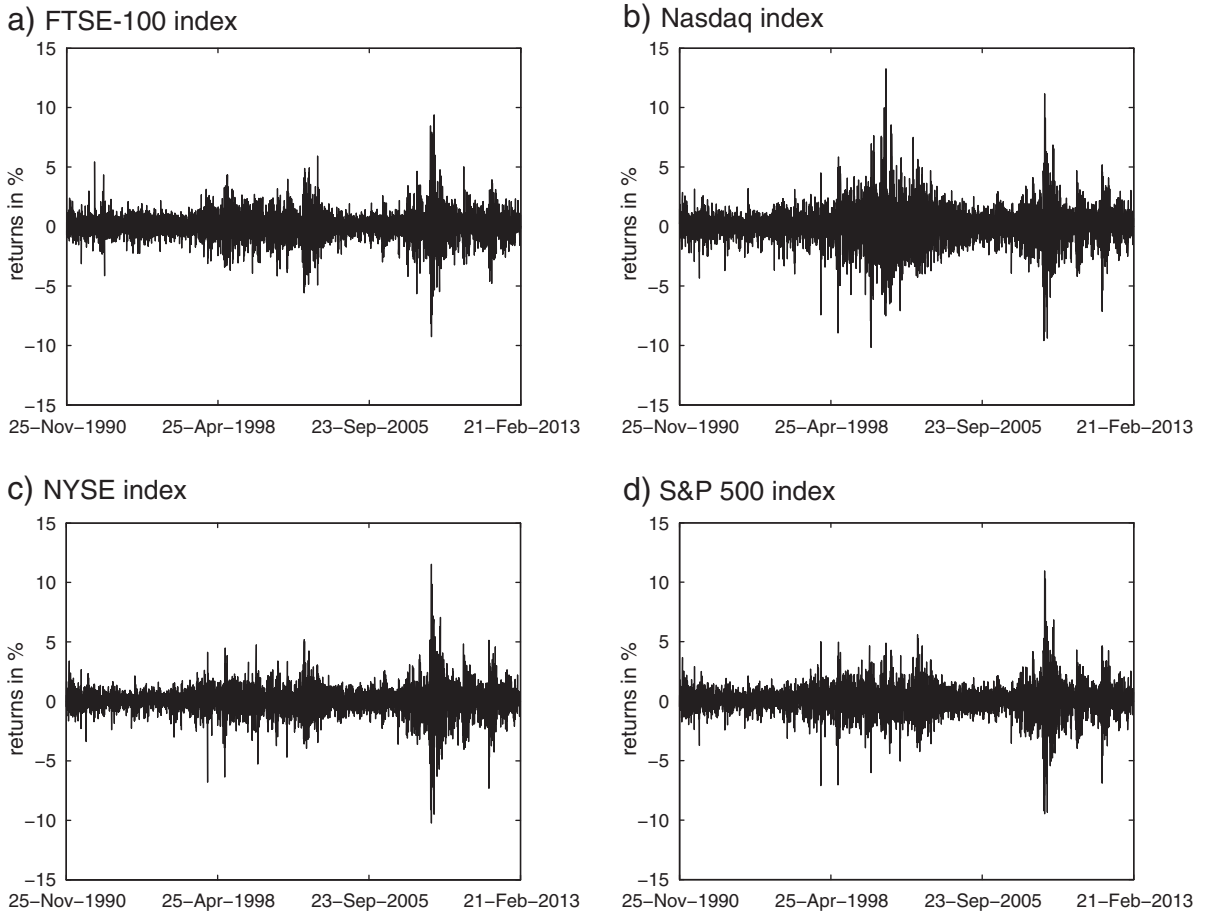


Fig. 5. Returns in percentage of several stock market indexes.

Table 2
Descriptive statistics for stock index returns and Engle's ARCH test.

Stock index returns	FTSE-100	Nasdaq	NYSE	S&P 500
Mean	0.0194	0.0396	0.0285	0.0281
Variance	1.3209	2.4127	1.3103	1.3807
Skewness	0.1115	0.0755	0.3854	0.2331
Kurtosis	9.0428	8.8006	13.8450	11.6672
Engle's ARCH test	306.6120 ^a	306.3044 ^a	279.0858 ^a	248.7533 ^a

^a Stands for the rejection of the null hypothesis at any relevant sig. level.

1991, corresponds to the first major market correction in the post crash of October 19, 1987 era and February 27, 2007 is the day when the big decline in Chinese stocks occurred and the news of the weakness in some key readings on the U.S. economy. Finally, October 27, 1997, is identified as outlier for NYSE index and it corresponds to a mini crash caused by an economic crisis in Asia.

5.2. Out of sample performance

For a full evaluation of the results, we perform out of sample conditional tests on the MCRRs calculated by using M0, approach [1] 'correcting outliers' (methods M1, M2) and approaches [2] 'complex distributions' and [3] 'robust estimation'.

By definition, the failure rate of a model is the number of times the estimated MCRRs are smaller than the returns (in absolute value). If the model is correctly specified, the failure rate should be equal to the pre specified MCRR level (in our case, 5%).⁶

⁶ The failure rate for a long position is obtained as the percentage of negative returns smaller than the one day ahead MCRRs calculated for long positions. Analogously, for a short position the failure rate is estimated as the percentage of positive returns larger than the one day ahead MCRRs calculated for short positions (see Giot and Laurent, 2003, 2004).

Table 3

Results of the test for sign and size biases (and standard errors).

	ϕ_0	ϕ_1	ϕ_2	n_3	nR^2
FTSE-100	1.064 (0.000)	0.082 (0.192)	7.282 (0.051)	13.932 (0.001)	19.744 ^a
Nasdaq	0.931 (0.000)	0.235 (0.000)	0.822 (0.779)	8.225 (0.009)	46.855 ^a
NYSE	0.988 (0.000)	0.146 (0.040)	2.144 (0.604)	14.697 (0.002)	29.243 ^a
S&P 500	0.984 (0.000)	0.154 (0.030)	2.002 (0.625)	14.196 (0.002)	32.071 ^a

^a Stands for the rejection of the null hypothesis of the joint test at any relevant significance level.**Table 4**

Days (mm/dd/yy) identified as possible outliers at 5% sig. level.

FTSE-100	Nasdaq	NYSE	S&P 500
–	11/15/91	11/15/91 10/27/97 02/27/07	11/15/91 02/27/07

Therefore, we calculate the MCRRs for one day horizon for both long and short positions and then check if these MCRRs are exceeded by price movements in day $t + 1$. We roll this process forward and calculate the MCRRs for 504 days, that is, we use the first 5097 observations for the estimation of the models, leaving the following 504 observations for the performance evaluation. This period spans from February 16, 2011 to February 19, 2013 and contains the last global financial crisis. In Table 5 we present the number of violations of the MCRR estimates fitting either a Gaussian GARCH(1,1) or a Gaussian GJR(1,1).

Since the calculation of the empirical failure rate defines a sequence of ones (MCRR violation) and zeros (no MCRR violation), we can test if the theoretical failure rate, f , is equal to 5%, i.e., $H_0 : f = 5\%$ vs. $H_1 : f \neq 5\%$. Standard evaluation of the failure rate proceeds by simply comparing the percentage of exceedances to the true failure rate. However, as pointed out in the works by West (1996) and McCracken (2000), when parameters are estimated, parameter uncertainty can play a role in out of sample inference. According to Christoffersen (1998), testing for conditional coverage is important in the presence of higher order dynamics and this author proposed a procedure that is composed of three tests. The first tests for the unconditional coverage (denoted LR_{uc}) and it is a standard likelihood ratio test are given by

$$LR_{uc} = -2\log[L(p; I_1, I_2, \dots, I_n)/L(\hat{\pi}; I_1, I_2, \dots, I_n)] \underset{\sim}{asy} \chi^2(1),$$

where $\{I_t\}_{t=1}^n$ is the indicator sequence, p is the theoretical coverage, $\hat{\pi} = n_1/(n_0 + n_1)$ is the maximum likelihood estimate of the alternative failure rate π , n_0 is the number of zeros and n_1 is the number of ones in the sequence $\{I_t\}_{t=1}^n$.

The second tests for the independence part of the conditional coverage hypothesis (denoted LR_{ind}) and it is also a likelihood ratio test

$$LR_{ind} = -2\log\left[L\left(\hat{\Pi}_2; I_1, I_2, \dots, I_n\right)/L\left(\hat{\Pi}_1; I_1, I_2, \dots, I_n\right)\right] \underset{\sim}{asy} \chi^2(1),$$

where

$$\hat{\Pi}_1 = \begin{pmatrix} n_{00}/(n_{00} + n_{01}) & n_{01}/(n_{00} + n_{01}) \\ n_{10}/(n_{10} + n_{11}) & n_{11}/(n_{10} + n_{11}) \end{pmatrix}, \quad \hat{\Pi}_2 = \begin{pmatrix} 1 - \hat{\pi}_2 \hat{\pi}_2 & \hat{\pi}_2 \hat{\pi}_2 \\ 1 - \hat{\pi}_2 \hat{\pi}_2 & \hat{\pi}_2 \hat{\pi}_2 \end{pmatrix},$$

n_{ij} is the number of observations with value i followed by j and $\hat{\pi}_2 = (n_{01} + n_{11})/(n_{00} + n_{10} + n_{01} + n_{11})$. Finally, the third is a joint test of coverage and independence (denoted LR_{cc}) given by:

$$LR_{cc} = -2\log\left[L(p; I_1, I_2, \dots, I_n)/L(\hat{\Pi}_1; I_1, I_2, \dots, I_n)\right] \underset{\sim}{asy} \chi^2(1).$$

With this complete procedure it is possible to check if the dynamics or the error distribution is misspecified or both. In Table 6 we report the results of the likelihood ratio tests for conditional coverage.

From Table 5 we can conclude that, for the four considered series and long position, the MCRR number of violations (in percentage) obtained from all methods always exceeds the 5% nominal value, indicating that these models tend to over reject. In general, for all positions and models, the closest failure rate to 5% is achieved by M5 method (which occurs in more than 60% of

Table 5

Estimates of the failure rate (proportions of exceedances) obtained one day ahead. The MCRRs are computed to cover 95% of expected losses.

	FTSE-100		Nasdaq		NYSE		S&P 500	
	Long P.	Short P.	Long P.	Short P.	Long P.	Short P.	Long P.	Short P.
<i>Estimating MCRRs with GARCH(1,1)</i>								
M0	7.3%	5.4%	5.6%	4.4%	6.7%	3.8%	6.3%	4.8%
M1	7.3%	5.4%	5.6%	4.4%	6.7%	3.6%	6.3%	4.8%
M2	7.3%	5.4%	5.6%	4.4%	6.7%	3.6%	6.5%	4.8%
M4	6.7%	5.6%	5.8%	4.6%	6.7%	3.8%	6.5%	5.4%
M5	6.7%	5.6%	5.4%	6.0%	6.7%	4.8%	6.0%	5.0%
M6	7.1%	5.2%	5.8%	4.5%	6.7%	3.4%	5.8%	4.8%
<i>Estimating MCRRs with GJR(1,1)</i>								
M0	6.3%	4.2%	5.4%	4.2%	6.7%	3.4%	6.3%	3.8%
M1	6.3%	4.2%	5.6%	4.4%	6.7%	3.4%	6.1%	4.0%
M2	6.3%	4.2%	5.6%	4.5%	6.7%	3.4%	6.3%	4.0%
M4	6.9%	4.4%	5.6%	4.4%	6.9%	3.4%	6.2%	4.4%
M5	5.6%	4.8%	5.4%	4.8%	6.7%	4.2%	5.4%	5.0%

the cases). From Table 6 we see that all the methods pass the joint test of coverage and independence for Nasdaq, NYSE and S&P 500, which lead us to conclude that all methods are statistically equivalent in estimating the MCRRs (for a 95% coverage). The statistical equivalence of such different methods may be due to the fact that the out of sample period spans from February 16, 2011 to February 19, 2013, which includes the last global financial crisis and, therefore, it corresponds to a period of high volatility.

Table 6

p -values for the null hypotheses $f = 5\%$ of (Christoffersen, 1998)'s test.

Estimating MCRRs with GARCH(1,1):												
	FTSE-100			Nasdaq			NYSE			S&P 500		
	LR_{uc}	LR_{ind}	LR_{cc}	LR_{uc}	LR_{ind}	LR_{cc}	LR_{uc}	LR_{ind}	LR_{cc}	LR_{uc}	LR_{ind}	LR_{cc}
<i>Long position</i>												
M0	0.024	0.430	0.052	0.574	0.069	0.155	0.087	0.308	0.128	0.181	0.979	0.383
M1	0.024	0.430	0.052	0.574	0.069	0.155	0.087	0.308	0.128	0.181	0.979	0.383
M2	0.024	0.430	0.052	0.574	0.069	0.155	0.087	0.308	0.128	0.127	0.903	0.290
M4	0.024	0.430	0.053	0.448	0.554	0.593	0.087	0.308	0.128	0.127	0.903	0.290
M5	0.087	0.633	0.192	0.716	0.080	0.191	0.087	0.308	0.128	0.340	0.498	0.474
M6	0.087	0.308	0.128	0.448	0.060	0.120	0.087	0.308	0.128	0.448	0.554	0.593
<i>Short position</i>												
M0	0.967	0.106	0.257	0.504	0.968	0.765	0.186	0.189	0.170	0.805	0.444	0.689
M1	0.967	0.106	0.257	0.504	0.968	0.765	0.122	0.153	0.105	0.805	0.444	0.689
M2	0.967	0.106	0.257	0.504	0.968	0.765	0.122	0.153	0.105	0.805	0.444	0.689
M4	0.574	0.069	0.155	0.648	0.958	0.859	0.186	0.189	0.170	0.716	0.646	0.797
M5	0.574	0.069	0.155	0.340	0.498	0.474	0.805	0.444	0.689	0.716	0.646	0.797
M6	0.076	0.122	0.060	0.504	0.968	0.765	0.076	0.122	0.060	0.805	0.444	0.689
Estimating MCRRs with GJR(1,1):												
	FTSE-100			Nasdaq			NYSE			S&P 500		
	LR_{uc}	LR_{ind}	LR_{cc}	LR_{uc}	LR_{ind}	LR_{cc}	LR_{uc}	LR_{ind}	LR_{cc}	LR_{uc}	LR_{ind}	LR_{cc}
<i>Long position</i>												
M0	0.181	0.395	0.267	0.716	0.080	0.191	0.087	0.308	0.128	0.181	0.979	0.383
M1	0.181	0.395	0.267	0.574	0.069	0.155	0.087	0.308	0.128	0.181	0.979	0.383
M2	0.181	0.395	0.267	0.574	0.069	0.155	0.087	0.308	0.128	0.181	0.979	0.383
M4	0.058	0.269	0.084	0.574	0.614	0.710	0.058	0.269	0.084	0.252	0.945	0.485
M5	0.574	0.614	0.710	0.805	0.121	0.277	0.087	0.308	0.128	0.716	0.678	0.813
<i>Short position</i>												
M0	0.377	0.176	0.260	0.377	0.893	0.643	0.076	0.122	0.060	0.186	0.189	0.170
M1	0.377	0.176	0.260	0.504	0.968	0.765	0.076	0.122	0.060	0.271	0.231	0.255
M2	0.377	0.176	0.260	0.504	0.968	0.765	0.076	0.122	0.060	0.271	0.231	0.255
M4	0.504	0.156	0.280	0.504	0.968	0.765	0.076	0.122	0.060	0.504	0.328	0.474
M5	0.805	0.121	0.277	0.716	0.678	0.813	0.377	0.277	0.359	0.967	0.152	0.341

LR_{uc} , LR_{ind} , LR_{cc} stand for the LR test of unconditional coverage, the LR test of independence and the joint test of coverage and independence, respectively.

6. Conclusion

This paper shows the impact of outliers on the estimation of the MCRRs using a Gaussian GARCH(1,1) model and compares six different methods that attenuate these effects. The methods under evaluation cover three different approaches to deal with outliers: detection and correction, more complex models for accommodating outliers (specification or error distribution) and robust estimation. The first approach, which includes the proposal by Grané and Veiga (2010) with hard and soft thresholding, proceeds by detecting and correcting outliers before estimating these risk measures with GARCH type models. We also analyze the proposal by Franses and Ghijssels (1999) only for the GARCH(1,1) model. The second approach consists in fitting, respectively, t distributed or skewed t distributed GARCH type models directly to the data. Finally, the third approach considers the robust estimation by Muler and Yohai (2008).

The simulation results show that ignoring the presence of outliers produces biases on the MCRR estimates that are quite severe for small and moderate sample sizes, especially if the magnitude of the outlier is moderate or large. All methods studied in this work produce MCRR estimates that increase with the magnitude of the outlier, but tend to decrease with the sample size. Ignoring the presence of outliers with high magnitudes, makes outlier effects persist more, generating higher MCRR estimates even in large sample sizes. Additionally, we study the effect of model uncertainty (functional form and error distribution) on the estimation of MCRRs. Simulation results show that, in general, the biases on the MCRRs are expected to be higher under model uncertainty.

The empirical application and the out of sample results seem to indicate that either detecting and correcting for outliers, or using more complex models, or even robust estimation provides statistically equivalent failure rates in estimating the MCRRs (for 95% coverage). However, using both more general models and more complex distributions generate failure rates closer to 5%. The statistical equivalence of such different methods may be due to the fact that the out of sample period spans from February 16, 2011 to February 19, 2013, which corresponds to a high volatile period. Similar results in the context of stochastic volatility were found by Mao et al. (2013) when comparing the asymmetric autoregressive stochastic volatility model by Harvey and Shephard (1996) and more sophisticated models for approximately the same out of sample period.

Acknowledgements

Special thanks to N. Muler for providing the Matlab code to carry out the robust estimation of the GARCH model. This work has been partially supported by the Spanish Ministry of Science and Innovation (research projects MTM2010 17323, ECO2009 08100 and ECO2012 32401). We are grateful to two anonymous referees for their useful suggestions that have contributed to improve the quality of this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jempfin.2014.01.005>.

References

- Baillie, R., Bollerslev, T., 1989. The message in daily exchange rates: a conditional variance tale. *J. Bus. Econ. Stat.* 7, 297–309.
- Bilen, C., Huzurbazar, S., 2002. Wavelet-based detection of outliers in time series. *J. Comput. Graph. Stat.* 11, 311–327.
- Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *J. Econ.* 31, 307–327.
- Bollerslev, T., 1987. A conditionally heteroskedastic time series model for speculative prices and rates of return. *Rev. Econ. Stat.* 69, 542–547.
- Brooks, C., 2002. *Introductory Econometrics for Finance*. Cambridge, University Press.
- Brooks, C., Clare, A., Persaud, G., 2000. A word of caution on calculating market-based minimum capital risk requirements. *J. Bank. Financ.* 24, 1557–1574.
- Carnero, M., Peña, D., Ruiz, E., 2007. Effects of outliers on the identification and estimation of GARCH models. *J. Time Ser. Anal.* 28, 471–497.
- Carnero, M.A., Peña, D., Ruiz, E., 2008. Estimating and forecasting GARCH volatility in the presence of outliers. Working Paper AD-13, Instituto Valenciano de Investigaciones Económicas.
- Carnero, M., Peña, D., Ruiz, E., 2012. Estimating GARCH volatility in the presence of outliers. *Econ. Lett.* 114, 86–90.
- Charles, A., 2008. Forecasting volatility with outliers in GARCH models. *J. Forecast.* 27, 551–565.
- Charles, A., Darné, O., 2005. Outliers and GARCH models in financial data. *Econ. Lett.* 86, 347–352.
- Chen, C., Liu, L., 1993a. Forecasting time series with outliers. *J. Forecast.* 12, 13–35.
- Chen, C., Liu, L., 1993b. Joint estimation of model parameters and outlier effects. *J. Am. Stat. Assoc.* 88, 284–297.
- Christoffersen, P., 1998. Evaluating interval forecasts. *Int. Econ. Rev.* 39, 841–862.
- Doornik, J., Ooms, M., 2005. Outlier detection in GARCH models. Discussion Paper, Tinbergen Institute.
- Droge, B., 2006. Minimax regret comparison of hard and soft thresholding for estimating a bounded mean. *Stat. Probab. Lett.* 76, 83–92.
- Duan, J.-C., Gauthier, G., Simonato, J.-G., Sasseville, C., 2006. Approximating the GJR-GARCH and EGARCH option pricing models analytically. *J. Comput. Financ.* 9, 41–69.
- Engle, R., Ng, V., 1993. Measuring and testing the impact of news on volatility. *J. Financ.* 48, 1749–1778.
- Fox, A., 1972. Outliers in time series. *J. R. Stat. Soc. B* 34, 350–363.
- Franses, P., Ghijssels, H., 1999. Additive outliers, GARCH and forecasting volatility. *Int. J. Forecast.* 15, 1–9.
- Franses, P., van Dijk, D., 2002. *Nonlinear Time Series Models in Empirical Finance*. Cambridge University Press, Cambridge, UK.
- Giot, P., Laurent, S., 2003. Value-at-risk for long and short trading positions. *J. Appl. Econ.* 18, 641–663.
- Giot, P., Laurent, S., 2004. Modelling daily value-at-risk using realized volatility and ARCH type models. *J. Empir. Financ.* 11, 379–398.
- Glosten, L., Jagannathan, R., Runkle, D., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *J. Financ.* 48, 1779–1801.
- Grané, A., Veiga, H., 2010. Wavelet-based detection of outliers in financial time series. *Comput. Stat. Data Anal.* 54, 2580–2593.
- Grossi, L., Laurini, F., 2009. A robust forward weighted lagrange multiplier test for conditional heteroscedasticity. *Comput. Stat. Data Anal.* 53, 2251–2263.
- Harvey, A., Shephard, N., 1996. Estimation of an asymmetric stochastic volatility model for asset returns. *J. Bus. Econ. Stat.* 14, 429–434.
- Hotta, L., Tsay, R., 1998. Outliers in GARCH processes. Manuscript Graduate School of Business, University of Chicago.

- Hsieh, D., 1993. Implications of nonlinear dynamics for financial risk management. *J. Financ. Quant. Anal.* 28, 41–64.
- Huber, P.J., 1981. *Robust Statistics*. Wiley, New York.
- Laurent, S., Peters, J., 2006. GARCH 4.2, Estimating and Forecasting ARCH Models. Timberlake Consultants.
- Ledolter, J., 1989. The effect of additive outliers on the forecasts from ARIMA models. *Int. J. Forecast.* 5, 231–240.
- Ling, S., McAleer, M., 2002. Stationarity and the existence of moments of a family of GARCH processes. *J. Econ.* 106, 109–117.
- Mao, X., Ruiz, E., Veiga, H., 2013. One for all: nesting asymmetric stochastic volatility models. Technical Report WP 13–11. Universidad Carlos III de Madrid.
- McCracken, M., 2000. Robust out-of-sample inference. *J. Econ.* 99, 195–223.
- Muler, N., Yohai, V.J., 2008. Robust estimates for GARCH models. *J. Stat. Plan. Infer.* 138, 2918–2940.
- Park, B.-J., 2002. An outlier robust GARCH model and forecasting volatility of exchange rate returns. *J. Forecast.* 21, 381–393.
- Peña, D., 2001. Outliers, influential observations and missing data. In: Peña, D., Tiao, G., Tsay, R. (Eds.), *A Course in Time Series*, New York. Wiley, pp. 136–170.
- Percival, D., Walden, A., 2000. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, New York.
- Sakata, S., White, H., 1998. High breakdown point conditional dispersion estimation with application to S&P500 daily returns volatility. *Econometrica* 66, 529–567.
- Taylor, S., 1986. *Modelling Financial Time Series*. Wiley, New York.
- Teräsvirta, T., 1996. Two stylized facts and the GARCH(1,1) model. Technical Report 96. Stockholm School of Economics.
- Van Dijk, D., Franses, P., Lucas, A., 1999. Testing for ARCH in the presence of additive outliers. *J. Appl. Econ.* 14, 539–562.
- Verhoeven, P., McAleer, M., 2000. Modelling outliers and extreme observations for ARMA-GARCH processes. Working paper. University of Western Australia.
- West, K., 1996. Asymptotic inference about predictive ability. *Econometrica* 64, 1067–1084.
- Zhang, X., 2004. Assessment of local influence in GARCH processes. *J. Time Ser. Anal.* 25 (2), 301–313.